

Mixing times and moving targets

Perla Sousi*

Peter Winkler†

October 19, 2012

Abstract

We consider irreducible Markov chains on a finite state space. We show that the mixing time of any such chain is equivalent to the maximum, over initial states x and moving large sets $(A_s)_s$, of the hitting time of $(A_s)_s$ starting from x . We prove that in the case of the d -dimensional torus the maximum hitting time of moving targets is equal to the maximum hitting time of stationary targets. Nevertheless, we construct a transitive graph where these two quantities are not equal, resolving an open question of Aldous and Fill on a “cat and mouse” game.

Keywords and phrases. Markov chain, mixing time, hitting time, rearrangement inequality.
MSC 2010 subject classifications. Primary 60J10; Secondary 60J27, 60G40.

1 Introduction

Mixing times and hitting times are fundamental notions for finite-state Markov chains. Both have been widely studied (see, e.g., [1] or [5] for background and numerous references) and a great variety of techniques have been developed to analyze them.

We begin by fixing some notation and reviewing previous work relating these two quantities.

Let $(X_t)_{t \geq 0}$ be an irreducible Markov chain on a finite state space with transition matrix P and stationary distribution π . For x, y in the state space we write

$$P^t(x, y) = \mathbb{P}_x(X_t = y),$$

for the transition probability in t steps.

Let $d(t) = \max_x \|P^t(x, \cdot) - \pi\|$, where $\|\mu - \nu\|$ stands for the total variation distance between the two probability measures μ and ν . The total variation mixing is defined as follows:

$$t_{\text{mix}}(\varepsilon) = \min\{t \geq 0 : d(t) \leq \varepsilon\}.$$

We use the convention that $t_{\text{mix}} = t_{\text{mix}}(1/4)$.

Before stating our first theorem, we introduce the maximum hitting time of “big” sets. Let $\alpha < 1/2$, then we define

$$t_{\text{H}}(\alpha) = \max_{x, A: \pi(A) \geq \alpha} \mathbb{E}_x[\tau_A],$$

*University of Cambridge, Cambridge, UK; p.sousi@statslab.cam.ac.uk

†Dartmouth College, Hanover, NH; peter.winkler@dartmouth.edu

where τ_A stands for the first hitting time of the set A .

We say that two real-valued functions f and g are *equivalent*, denoted $f \asymp g$, if there are universal positive constants c and c' such that $cf \leq g \leq c'f$. If the constants are allowed to depend on a parameter α , we write $f \asymp_\alpha g$.

Aldous (1981) related mixing and hitting times by proving that $t_{\text{cts}} \asymp \max_{x,A} \pi(A) \mathbb{E}_x[\tau_A]$ for all reversible chains, where t_{cts} is the mixing time of the continuous time chain. In two independent recent papers by Imbuzeiro Oliveira [4] and Peres and Sousi [8] it was proved that for all reversible chains, if $\alpha < 1/2$, then

$$t_L \asymp_\alpha t_H(\alpha), \quad (1.1)$$

where t_L is the mixing time of the lazy version of the chain, i.e. the chain with transition matrix $\frac{P+I}{2}$.

Very recently, Griffiths, Kang, Imbuzeiro Oliveira and Patel [3] showed that $t_H(\alpha) \leq t_H(1/2)/\alpha$ for all $\alpha < 1/2$. Hence this together with (1.1) or with Aldous' result implies that for all reversible chains if $\alpha \leq 1/2$, then

$$t_L \asymp_\alpha t_H(\alpha),$$

with the equivalence failing if $\alpha > 1/2$.

For non-reversible chains equation (1.1) may fail, e.g. for biased random walk on the cycle \mathbb{Z}_n we have $t_L \asymp n^2$, while $t_H(\alpha) \asymp n$, for any $\alpha > 0$. During a lecture on [8] by Yuval Peres, Guy Kindler proposed that for non-reversible chains the right analogue of (1.1) involves moving targets. Our first result establishes this equivalence.

Let $\alpha \in (0, 1)$ and $\mathcal{A}(\alpha)$ denote the collection of sequences of sets defined as follows:

$$\mathcal{A}(\alpha) = \{A = (A_t)_{t \geq 0} : \forall t \geq 0, \pi(A_t) \geq \alpha\}.$$

For $A \in \mathcal{A}(\alpha)$ define $\tau_A = \inf\{t \geq 0 : X_t \in A_t\}$ and

$$t_{\text{mov}}(\alpha) = \sup_{x, A \in \mathcal{A}(\alpha)} \mathbb{E}_x[\tau_A].$$

Theorem 1.1. *For $\alpha < 1/2$, $t_{\text{mix}} \asymp t_{\text{mov}}(\alpha)$.*

We will prove Theorem 1.1 in Section 2.

Remark 1.2. We note that Theorem 1.1 does not require the chain to be either lazy or reversible, as is the case for (1.1). In this setting the equivalence holds for any chain.

Theorem 1.1 and (1.1) immediately give that for all reversible lazy chains and for any $\alpha < 1/2$

$$t_{\text{mov}}(\alpha) \asymp_\alpha t_H(\alpha).$$

If the chain is not reversible, though, the above equivalence can fail. For instance, for the biased random walk on \mathbb{Z}_n , if $A = (A_i)_i$ are sets moving at the same speed as the random walk, then $\mathbb{E}[\tau_A] \asymp n^2$ agreeing with the mixing time t_L .

We next consider the problem of colliding with a moving target on a graph. In the following theorem we show that in the case of toroidal grids, the best strategy for the target, to avoid collision as long as possible, is to stay in place at the maximum distance from the starting point. As a corollary, we show that in the 1-dimensional case the two quantities t_H and t_{mov} are equal.

Theorem 1.3. *Let X be a lazy simple random walk on \mathbb{Z}_n^d and $f : \mathbb{N} \rightarrow \mathbb{Z}_n^d$ a function. Then setting $a = (\lfloor n/2 \rfloor, \dots, \lfloor n/2 \rfloor)$ we have for all t*

$$\mathbb{P}_0(X_1 \neq f(1), \dots, X_t \neq f(t)) \leq \mathbb{P}_0(X_1 \neq a, \dots, X_t \neq a).$$

Remark 1.4. Note that if the random walk X on \mathbb{Z}_n^d is not lazy, then one can always choose a function $f : \mathbb{N} \rightarrow \mathbb{Z}_n^d$ so that

$$\mathbb{P}_0(X_1 \neq f(1), \dots, X_t \neq f(t)) = 1,$$

and hence the conclusion of Theorem 1.3 fails.

Corollary 1.5. *Let X be a lazy simple random walk on $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$. Then for all n, α we have*

$$t_H(\alpha) = t_{\text{mov}}(\alpha).$$

We prove Theorem 1.3 and Corollary 1.5 in Section 3 using a discrete version of rearrangement inequalities. We employ a polarization technique which has been used extensively in the continuous setting to prove several classical rearrangement inequalities (see, for instance, [2]). As a by-product of the discrete rearrangement inequality, we also prove that the expected volume of the “sausage” around a discrete lazy simple random walk on \mathbb{Z}^d with drift is minimized when the drift is equal to 0.

Proposition 1.6. *Let X be a lazy simple random walk on \mathbb{Z}^d and let $f : \mathbb{N} \rightarrow \mathbb{Z}^d$ be a function. Then for all $t \in \mathbb{N}$ and all $n \in \mathbb{N}$*

$$\mathbb{E} \left[\text{vol} \left(\bigcup_{s=0}^t (X_s + f(s) + Q_n) \right) \right] \geq \mathbb{E} \left[\text{vol} \left(\bigcup_{s=0}^t (X_s + Q_n) \right) \right],$$

where $Q_n = [-n, n]^d$.

A more general isoperimetric inequality for the expected volume of the Wiener sausage has been proved in [7]; the stronger Proposition 1.6 makes use of the symmetries of \mathbb{Z}^d and does not hold in general.

Finally, in the last theorem, we show that the equality of Proposition 1.5 is not always true for a reversible Markov chain. This resolves an open question of Aldous [1, Chapter 4, Open Problem 20] and of Imbuzeiro Oliveira [6].

We say that X is a continuous time random walk on a graph if it stays at every vertex for an exponential amount of time of mean 1, and then jumps to one of the neighbours uniformly at random.

Theorem 1.7. *There exists a transitive graph $G = (V, E)$ such that if X is a continuous time or lazy random walk on G , then*

$$\max_{x, y} \mathbb{E}_x[\tau_y] < \sup_{x, f \in V^{\mathbb{R}_+}} \mathbb{E}_x[\tau_f],$$

where $\tau_f = \inf\{t \geq 0 : X_t = f(t)\}$.

In [1] and [6] this was stated as a cat and mouse problem and it was conjectured that the best strategy for the mouse to maximize the expected capture time is to stay in place. In our graph G we show that this is not the case. We prove Theorem 1.7 in Section 4.

2 Moving targets

Proof of Theorem 1.1. We first show that $t_{\text{mov}} \leq c_1 t_{\text{mix}}$, where c_1 is a positive constant.

Let $t = t_{\text{mix}}(\alpha/2) \leq \lceil \log_2(1/\alpha) \rceil t_{\text{mix}}$. Then for all x and all sets A we have

$$P^t(x, A) \geq \pi(A) - \frac{\alpha}{2}.$$

Take a sequence of sets $A = (A_s) \in \mathcal{A}(\alpha)$. Then for all s and all starting points x we have

$$P^t(x, A_s) \geq \frac{\alpha}{2}. \quad (2.1)$$

If $\tau = \min\{k \geq 0 : X_{kt} \in A_{kt}\}$, then obviously we have $\tau_A \leq t\tau$. By (2.1), it follows that τ is stochastically dominated by a geometric random variable of success probability $\alpha/2$. Therefore,

$$\mathbb{E}_x[\tau_A] \leq t\mathbb{E}_x[\tau] \leq \frac{2t}{\alpha},$$

and hence this gives that

$$t_{\text{mov}} \leq \frac{2\lceil \log_2(1/\alpha) \rceil}{\alpha} t_{\text{mix}}$$

and this completes the proof of the upper bound.

We now show the other direction, i.e. that there exists a positive constant c_2 so that

$$t_{\text{mix}} \leq c_2 t_{\text{mov}}(\alpha).$$

Since $\alpha < 1/2$, there exists $\varepsilon > 0$ such that $\alpha + \varepsilon < 1/2$. By [5, 4.35], it follows that there exists a positive constant c_3 such that $t_{\text{mix}}(\alpha + \varepsilon) \geq c_3 t_{\text{mix}}$. Let $t < t_{\text{mix}}(\alpha + \varepsilon)$. Then this means that there exists x and a set A so that

$$P^t(x, A) < \pi(A) - (\alpha + \varepsilon). \quad (2.2)$$

From that we immediately get that $\pi(A) > \alpha + \varepsilon$. We now use the set A to define a sequence of sets (B_s) as follows: for $s < t$ define

$$B_s = \{y : P^{t-s}(y, A) > \pi(A) - \alpha\}$$

and for $s \geq t$ we let $B_s = \Omega$. Since π is stationary, it follows that

$$\pi(A) = \sum_{y \in B_s} P^{t-s}(y, A)\pi(y) + \sum_{y \in B_s^c} P^{t-s}(y, A)\pi(y) \leq \pi(B_s) + \pi(A) - \alpha.$$

Rearranging, gives that $\pi(B_s) \geq \alpha$ for all s . We write $\tau_B = \min\{t \geq 0 : X_t \in B_t\}$. We will show for a constant θ to be determined later we have that

$$\mathbb{E}_x[\tau_B] \geq \theta t. \quad (2.3)$$

We will show that for a θ to be specified later, assuming

$$\max_z \mathbb{E}_z[\tau_B] \leq \theta t \quad (2.4)$$

will yield a contradiction.

By Markov's inequality and (2.4) we have that for all z

$$\mathbb{P}_z(\tau_B \leq t) \geq 1 - \theta.$$

By the strong Markov property applied to the stopping time τ_B and Markov's inequality we have

$$\mathbb{P}_x(X_t \in A) \geq \mathbb{P}_x(X_t \in A \mid \tau_B \leq t) \mathbb{P}_x(\tau_B \leq t) \geq \inf_{s \leq t} \inf_{w \in B_s} \mathbb{P}_w(X_{t-s} \in A) (1 - \theta) \geq (\pi(A) - \alpha)(1 - \theta),$$

which by choosing θ small enough can be made bigger than $\pi(A) - (\alpha + \varepsilon)$. This contradicts the choice of x in (2.2). Therefore (2.3) holds and this completes the proof. \square

3 Collision with a moving target on \mathbb{Z}_n^d and \mathbb{Z}^d

In this section we prove Theorem 1.3 and Proposition 1.6. We start by introducing some notation and background on rearrangement inequalities. We follow closely Section 2.1 of Burchard and Schmuckenschläger [2].

3.1 Notation and background

Let M be a metric space. A *reflection* $\sigma : M \rightarrow M$ is an isometry such that

- $\sigma^2 x = x$ for all $x \in M$;
- M is the disjoint union of the set of fixed points H^0 , and two half spaces H^- and H^+ which are exchanged by σ , i.e.

$$\begin{aligned}\sigma x &= x & x \in H^0, \\ \sigma H^+ &= H^-;\end{aligned}$$

- $d(x, y) < d(x, \sigma y)$ for all $x, y \in H^+$.

From now on whenever we define a reflection σ we will specify H^+ and H^- .

The *two-point rearrangement* of a function f is defined to be

$$f^\sigma(x) = \begin{cases} \max\{f(x), f(\sigma x)\}, & \text{if } x \in H^+; \\ \min\{f(x), f(\sigma x)\}, & \text{if } x \in H^-; \\ f(x), & \text{if } x \in H^0. \end{cases}$$

By taking $f = \mathbf{1}(A)$ we get that the two-point rearrangement of a set A , denoted A^σ , satisfies

$$\begin{aligned}A^\sigma \cap H^+ &= (A \cup \sigma A) \cap H^+ \\ A^\sigma \cap H^- &= (A \cap \sigma A) \cap H^-.\end{aligned}$$

We now recall a combinatorial lemma from [2, Lemma 2.6].

Consider the two-point space $\{+, -\}$ with the metric defined by $d(+, -) = 1$. The map σ that exchanges $+$ and $-$ is a reflection with no fixed points and with $H^+ = \{+\}$ and $H^- = \{-\}$ as the positive and negative half-spaces. For any function φ on $\{+, -\}$, let φ^σ be the corresponding two-point rearrangement of φ :

$$\varphi^\sigma(+) = \max\{\varphi(+), \varphi(-)\} \quad \text{and} \quad \varphi^\sigma(-) = \min\{\varphi(+), \varphi(-)\}. \quad (3.1)$$

Lemma 3.1 (Burchard and Schmuckenschläger [2]). *Let $\varphi_1, \dots, \varphi_n$ be nonnegative functions on the set $\{+, -\}$. For each pair ij , let $k_{i,j}(\varepsilon, \varepsilon') = a_{ij} + b_{ij}\mathbf{1}(\varepsilon = \varepsilon')$ with $a_{ij}, b_{ij} \geq 0$. Consider the function*

$$J(\varphi_1, \dots, \varphi_n) = \sum_{\pm} \prod_{1 \leq i \leq n} \varphi_i(\varepsilon_i) \prod_{1 \leq i \leq j \leq n} k_{i,j}(\varepsilon_i, \varepsilon_j).$$

Then

$$J(\varphi_1, \dots, \varphi_n) \leq J(\varphi_1^\sigma, \dots, \varphi_n^\sigma).$$

3.2 Random walk on \mathbb{Z}_n^d

Lemma 3.2. *Let σ be a reflection in \mathbb{Z}_n^d and X a lazy simple random walk in \mathbb{Z}_n^d . Then for all times t , all starting states b and all sets $D_i \subseteq \mathbb{Z}_n^d$ we have*

$$\mathbb{P}(X_1 \in D_1, \dots, X_t \in D_t \mid X_0 \in \{b\}) \leq \mathbb{P}(X_1 \in D_1^\sigma, \dots, X_t \in D_t^\sigma \mid X_0 \in \{b\}^\sigma).$$

Proof. Let $p(x, y)$ be the transition probability in one step of the lazy simple random walk in \mathbb{Z}_n^d , i.e.

$$p(x, y) = \mathbf{1}(x = y) \frac{1}{2} + \mathbf{1}(|x - y| = 1) \frac{1}{4d}.$$

By the Markov property we have

$$\mathbb{P}(X_1 \in D_1, \dots, X_t \in D_t \mid X_0 \in \{b\}) = \sum_{x_0, \dots, x_t} \prod_{i=1}^t p(x_{i-1}, x_i) \prod_{i=0}^t \mathbf{1}(x_i \in D_i),$$

where $D_0 = \{b\}$. If H^+ and H^- are the positive and negative respectively half spaces exchanged by σ , then we can write the above sum

$$\sum_{x_0, \dots, x_t} \prod_{i=1}^t p(x_{i-1}, x_i) \prod_{i=0}^t \mathbf{1}(x_i \in D_i) = \sum_{x_0, \dots, x_t \in H^+} \sum_{\pm} \prod_{i=1}^t p(x_{i-1}^\pm, x_i^\pm) \prod_{i=0}^t \mathbf{1}(x_i^\pm \in D_i),$$

where

$$x^+ = \begin{cases} x, & \text{if } x \in H^+ \\ \sigma x, & \text{if } x \in H^- \end{cases} \quad \text{and} \quad x^- = \begin{cases} \sigma x, & \text{if } x \in H^+ \\ x, & \text{if } x \in H^- \end{cases}. \quad (3.2)$$

We now fix a choice of $x_1, \dots, x_t \in H^+$. It suffices to show that

$$\sum_{\pm} \prod_{i=1}^t p(x_{i-1}^\pm, x_i^\pm) \prod_{i=0}^t \mathbf{1}(x_i^\pm \in D_i) \leq \sum_{\pm} \prod_{i=1}^t p(x_{i-1}^\pm, x_i^\pm) \prod_{i=0}^t \mathbf{1}(x_i^\pm \in D_i^\sigma). \quad (3.3)$$

For $\varepsilon, \varepsilon' \in \{+, -\}$ we define $k_{i,j}(\varepsilon, \varepsilon') = 1$ if $j - i \neq 1$ and otherwise

$$k_{i-1,i}(\varepsilon, \varepsilon') = p(x_{i-1}^-, x_i^+) + \mathbf{1}(\varepsilon = \varepsilon')(p(x_{i-1}^+, x_i^+) - p(x_{i-1}^-, x_i^+)).$$

By the definition of the transition probability we have $p(x_{i-1}^-, x_i^+) \leq p(x_{i-1}^+, x_i^+)$ for $x_{i-1}, x_i \in H^+$. Therefore $k_{i,j}$ satisfies the assumptions of Lemma 3.1 and if we set $\varphi_i(\varepsilon) = \mathbf{1}(x_i^\varepsilon \in D_i)$, then we can write

$$\sum_{\pm} \prod_{i=1}^t p(x_{i-1}^\pm, x_i^\pm) \prod_{i=0}^t \mathbf{1}(x_i^\pm \in D_i) = \sum_{\pm} \prod_{i=0}^t \varphi_i(\varepsilon_i) \prod_{0 \leq i < j \leq t} k_{i,j}(\varepsilon_i, \varepsilon_j). \quad (3.4)$$

Applying Lemma 3.1 we infer

$$\sum_{\pm} \prod_{i=0}^t \varphi_i(\varepsilon_i) \prod_{0 \leq i < j \leq t} k_{i,j}(\varepsilon_i, \varepsilon_j) \leq \sum_{\pm} \prod_{i=0}^t \varphi_i^\sigma(\varepsilon_i) \prod_{0 \leq i < j \leq t} k_{i,j}(\varepsilon_i, \varepsilon_j). \quad (3.5)$$

Since $\varphi_i^\sigma(\varepsilon) = \mathbf{1}(x_i^\varepsilon \in D_i^\sigma)$, inequality (3.5) together with (3.4) concludes the proof of (3.3) and thus completes the proof of the lemma. \square

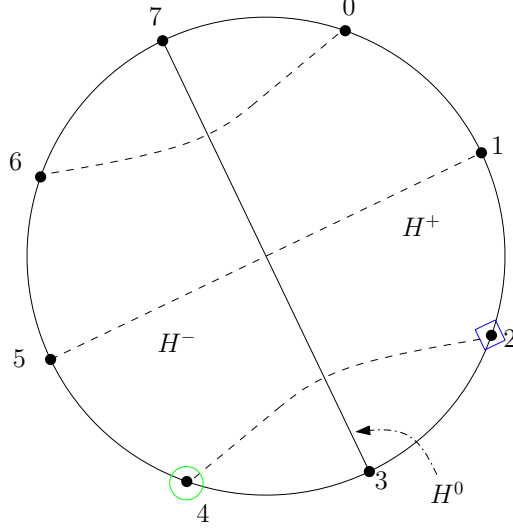


Figure 1: A reflection on \mathbb{Z}_8

Remark 3.3. Note that it is essential that the random walk on \mathbb{Z}_n^d be lazy. In the proof above this was used to show that the kernel k satisfies the assumptions of Lemma 3.1.

Proof of Theorem 1.3. We first prove the theorem for $d = 1$.

For $i = 1, \dots, t$ we write $D_i = \mathbb{Z}_n \setminus \{f(i)\}$. Then we have

$$\mathbb{P}_0(X_1 \neq f(1), \dots, X_t \neq f(t)) = \mathbb{P}_0(X_1 \in D_1, \dots, X_t \in D_t).$$

We now want to find a sequence of reflections $\sigma_1, \dots, \sigma_k$ such that $D_i^{\sigma_1 \dots \sigma_k} = \mathbb{Z}_n \setminus \{a\}$.

We first give the reflection σ such that $D_1^\sigma = \mathbb{Z}_n \setminus \{a\}$. We carry out all the details in the case when n is odd and $f(1) + a$ is even and satisfies $f(1) + a \geq n - 1$. The other cases follow similarly. We define

$$\sigma_1(x) = (a + f(1) - x) \bmod n$$

and we let

$$H^+ = \mathbb{Z}_n \cap \left(\left(\frac{a + f(1)}{2}, n - 1 \right] \cup \left[0, \frac{a + f(1)}{2} - \frac{n - 1}{2} \right) \right) \quad \text{and}$$

$$H^- = (H^+)^c \setminus \left\{ \frac{a + f(1)}{2} \right\}.$$

Then with this definition of H^+ and H^- it is clear that $D_1^{\sigma_1} = \mathbb{Z}_n \setminus \{a\}$ and $(\mathbb{Z}_n \setminus \{a\})^{\sigma_1} = \mathbb{Z}_n \setminus \{a\}$ and $\{0\}^{\sigma_1} = \{0\}$.

Having symmetrized the set D_1 , we now want to find a reflection σ_2 such that $D_2^{\sigma_1 \sigma_2} = \mathbb{Z}_n \setminus \{a\}$. To do that we use exactly the same construction as for σ_1 above. Hence we get $(\mathbb{Z}_n \setminus \{a\})^{\sigma_2} = \mathbb{Z}_n \setminus \{a\}$ and $\{0\}^{\sigma_2} = \{0\}$. Therefore $D_1^{\sigma_1 \sigma_2} = \mathbb{Z}_n \setminus \{a\}$. Continuing in this manner we find $k \leq t$ reflections $\sigma_1, \dots, \sigma_k$ such that for all i

$$D_i^{\sigma_1 \dots \sigma_k} = \mathbb{Z}_n \setminus \{a\}.$$

Proof. Since $D_s = -D_s$ for all s , we have

$$\begin{aligned}\mathbb{E}\left[\text{vol}\left(\bigcup_{s=0}^t (X_s + D_s)\right)\right] &= \mathbb{E}\left[\sum_{x_0 \in \mathbb{Z}^d} \mathbf{1}\left(x_0 \in \bigcup_{s=0}^t (X_s + D_s)\right)\right] \\ &= \mathbb{E}\left[\sum_{x_0 \in \mathbb{Z}^d} \mathbf{1}(\exists s \leq t : X_s \in x_0 + D_s)\right].\end{aligned}$$

Let $p(x, y)$ be the transition probability in one step of the lazy simple random walk in \mathbb{Z}^d , i.e.

$$p(x, y) = \mathbf{1}(x = y) \frac{1}{2} + \mathbf{1}(|x - y| = 1) \frac{1}{4d}.$$

Then the Markov property of the random walk gives

$$\mathbb{P}(\exists s \leq t : X_s \in x_0 + D_s) = 1 - \sum_{y_1, \dots, y_t} \prod_{i=1}^t p(y_{i-1}, y_i) \prod_{i=0}^t \mathbf{1}(y_i \notin x_0 + D_i), \quad (3.6)$$

where $y_0 = 0$. Changing variables to $y_i - x_0$ and noticing that $p(0, y_1 + x_0) = p(-x_0, y_1)$ gives that the sum appearing in the right-hand side of (3.6) is equal to

$$\sum_{y_1, \dots, y_t} p(-x_0, y_1) \mathbf{1}(-x_0 \notin D_0) \prod_{i=2}^t p(y_{i-1}, y_i) \prod_{i=1}^t \mathbf{1}(y_i \notin D_i).$$

Putting everything together in the expression for the expected volume of $\cup_{s \leq t} (\xi(s) + D_s)$ and changing variables from $-x_0$ to x_0 we get

$$\mathbb{E}\left[\text{vol}\left(\bigcup_{s=0}^t Q_n(f(s) + \xi(s))\right)\right] = \sum_{x_0, x_1, \dots, x_t} \prod_{i=1}^t p(x_{i-1}, x_i) \left(1 - \prod_{i=0}^t \mathbf{1}(x_i \notin D_i)\right). \quad (3.7)$$

Decomposing the above sum into the positive and negative half spaces of σ , the right hand side of (3.7) can be written as

$$\sum_{x_0, x_1, \dots, x_t \in H^+} \sum_{\pm} \prod_{i=1}^t p(x_{i-1}^{\pm}, x_i^{\pm}) \left(1 - \prod_{i=0}^t \mathbf{1}(x_i^{\pm} \notin D_i)\right),$$

where x^+ and x^- are as defined in (3.2) in the proof of Lemma 3.2. Repeating the same arguments as in the proof of (3.3) in Lemma 3.2 we get

$$\sum_{\pm} \prod_{i=1}^t p(x_{i-1}^{\pm}, x_i^{\pm}) \left(1 - \prod_{i=0}^t \mathbf{1}(x_i^{\pm} \notin D_i)\right) \geq \sum_{\pm} \prod_{i=1}^t p(x_{i-1}^{\pm}, x_i^{\pm}) \left(1 - \prod_{i=0}^t \mathbf{1}(x_i^{\pm} \notin D_i^{\sigma})\right).$$

Hence, we conclude that

$$\mathbb{E}\left[\text{vol}\left(\bigcup_{s=0}^t (\xi(s) + D_s)\right)\right] \geq \mathbb{E}\left[\text{vol}\left(\bigcup_{s=0}^t (\xi(s) + D_s^{\sigma})\right)\right]$$

and this finishes the proof of the lemma. \square

Proof of Proposition 1.6. Let $r > 0, x \in \mathbb{Z}^d$ and $Q_r(x) = [-r+x_1, r+x_1] \times \dots \times [-r+x_d, r+x_d]$ be the box in \mathbb{Z}^d of side length $2r+1$ centered at x . We want to show that

$$\mathbb{E} \left[\text{vol} \left(\bigcup_{s=0}^t (\xi(s) + Q_n(f(s))) \right) \right] \geq \mathbb{E} \left[\text{vol} \left(\bigcup_{s=0}^t (\xi(s) + Q_n) \right) \right],$$

where $Q_n = [-n, n]^d$ as defined in the statement of the proposition.

We now want to find a sequence of reflections $\sigma_1, \dots, \sigma_k$ such that $Q_n(f(s))^{\sigma_1 \dots \sigma_k} = Q_n$ for all $s \leq t$.

First we show how to bring a non-centered interval to a centered one in \mathbb{Z} . Let $A = [a-n, a+n]$, where $a \in \mathbb{Z}$ and $n \in \mathbb{N}$. Define the reflection σ around the point $a/2$ via

$$\sigma(x) = a - x.$$

Then it is clear that σ maps the interval $[a-n, a+n]$ to the interval $[-n, n]$. If $a > 0$, define $H^+ = \{k \in \mathbb{Z} : k \leq a/2\}$ and H^- to be its complement. If $a < 0$, define $H^+ = \{k \in \mathbb{Z} : k \geq a/2\}$. It is then easy to see that $A^\sigma = [-n, n]$ and $[-n, n]^\sigma = [-n, n]$.

Next we define reflections in \mathbb{Z}^d . Let $A = [a_1 - n_1, a_1 + n_1] \times \dots \times [a_d - n_d, a_d + n_d]$ and for $i = 1, \dots, d$ let

$$\sigma_i(x_1, \dots, x_d) = (x_1, \dots, x_{i-1}, a_i - x_i, x_{i+1}, \dots, x_d).$$

Then $A^{\sigma_1 \dots \sigma_d} = [-n_1, n_1] \times \dots \times [-n_d, n_d]$ and if B is a centered rectangle, then $B^{\sigma_1 \dots \sigma_d} = B$.

This way we see that there exist $k \leq td$ reflections $\sigma_1, \dots, \sigma_k$ such that

$$Q_n(f(s))^{\sigma_1 \dots \sigma_k} = Q_n \quad \text{for all } s \leq t.$$

Applying Lemma 3.5 k times concludes the proof of the proposition. \square

4 Better to run than hide

In this section we give the proof of Theorem 1.7. We first define a class of graphs indexed by n, m and denoted $G_{n,m}$. For $n = 2$ and $m = 12$ the graph is illustrated in Figure 3. We then prove that $G_{2,12}$ is an example of a graph satisfying the statement of Theorem 1.7 for a lazy discrete time walk. We conclude the section by proving that $G_{7,20}$ is such that it is best for a target to move in order to avoid collision with a continuous time walk.

Definition 4.1. Let m be a multiple of 4 and $G_{n,m}$ a graph on n^2m vertices divided into m clusters. We think of the clusters as the nodes of \mathbb{Z}_m and so we number them $0, \dots, m-1$. We give coordinates to each element of every cluster. The elements of cluster i have coordinates $i(a, b)$, where $a, b \in \mathbb{Z}_n$. We put an edge between

- (1) all pairs $i(a, b), j(c, d)$ with $|i - j| = 1$;
- (2) all pairs $i(a, b), j(a, d)$ with $b \neq d$, i even and $j = (i + m/4) \bmod m$;
- (3) all pairs $i(a, b), j(c, b)$ with $a \neq c$, i even and $j = (i - m/4) \bmod m$;
- (4) all pairs $i(a, b), j(c, b)$ with $a \neq c$, i odd and $j = (i + m/4) \bmod m$;
- (5) all pairs $i(a, b), j(a, d)$ with $b \neq d$, i odd and $j = (i - m/4) \bmod m$.

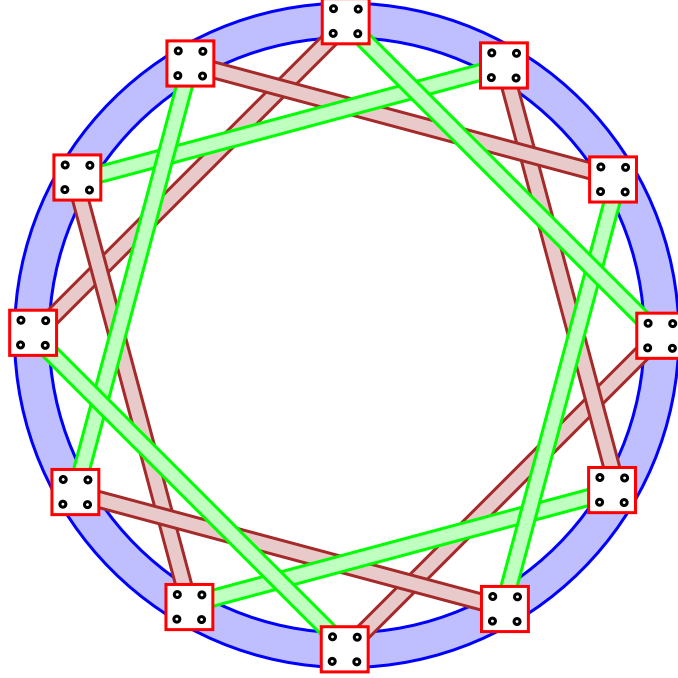


Figure 3: Graph $G_{2,12}$

We call the edges of type (1) “short” while the edges of type (2), (3), (4) and (5) “long”.

Remark 4.2. Intuitively, notice that for a fixed m as n goes to infinity, the long edges of $G_{n,m}$ are rarely used, and hence $G_{n,m}$ looks more like \mathbb{Z}_m .

Claim 4.1. $G_{n,m}$ is a vertex transitive graph.

Proof. Let $i(a, b), j(c, d)$ be two vertices of the graph $G_{n,m}$. In order to show that $G_{n,m}$ is vertex transitive, we need to construct an automorphism $\varphi : V \rightarrow V$ that preserves edges and satisfies $\varphi(i(a, b)) = j(c, d)$. We consider two separate cases, depending on whether $j - i$ is even or odd.

If $j - i$ is even, then we set

$$\varphi(k(u, v)) = ((k + j - i) \bmod m)((u + c - a) \bmod n, (v + d - b) \bmod n).$$

If $j - i$ is odd, then we set

$$\varphi(k(u, v)) = ((k + j - i) \bmod m)((v + c - b) \bmod n, (u + d - a) \bmod n).$$

It is straightforward to check that φ is an automorphism that preserves edges. \square

Lemma 4.3. Let X be a simple random walk on $G_{2,12}$ which is either discrete or continuous. Then we have

$$\max_{x,y} \mathbb{E}_x[\tau_y] = \mathbb{E}_{0(0,0)}[\tau_{6(1,1)}]. \quad (4.1)$$

Proof. It suffices to prove the lemma for a discrete time random walk. Since $G_{2,12}$ is vertex transitive, it follows that for $x \in V$ we have

$$\max_{x,y} \mathbb{E}_x[\tau_y] = \max_y \mathbb{E}_x[\tau_y].$$

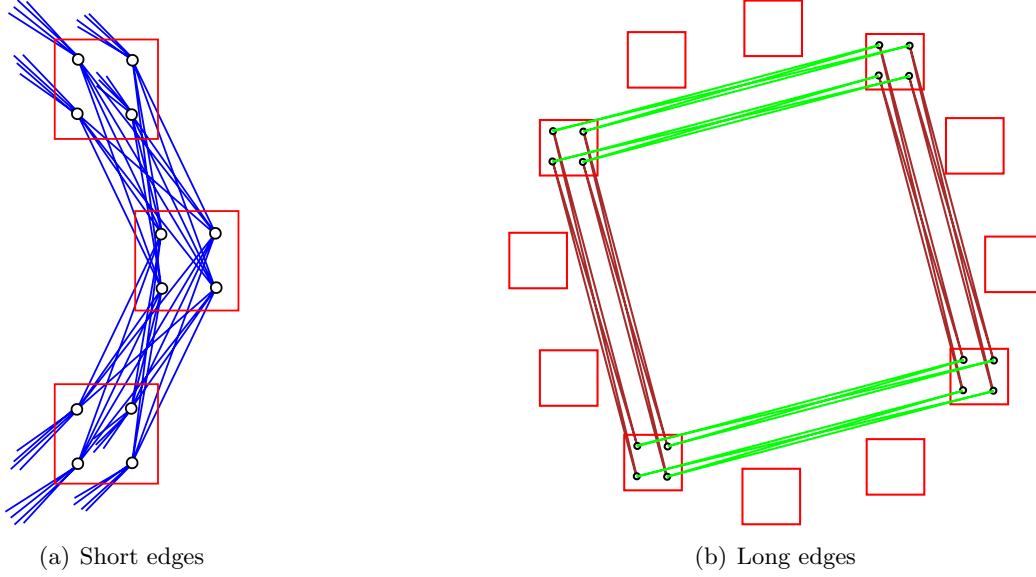


Figure 4: Edges of $G_{2,12}$

So taking $x = 0(0,0)$, it suffices to show that for all $a, b \in \mathbb{Z}_2$ we have

$$\mathbb{E}_{0(0,0)}[\tau_{6(a,b)}] = \max_y \mathbb{E}_{0(0,0)}[\tau_y]. \quad (4.2)$$

First we observe that starting from any point in cluster 0, the first time the random walk hits cluster 6, the position is uniform. Indeed, if we reach cluster 6 having used at least one short edge, then this is clear. If we use only long edges, then by the construction of the graph, with the first long edge we have randomized the column and with the second long edge we have randomized the row. Arguing similarly, if we start from cluster 0, the position at the first hitting time of cluster $i \neq 3, 9$ is uniform. Hence, if T_i is the first time that we hit cluster $i \neq 3, 9$, then

$$\mathbb{E}_{0(0,0)}[\tau_{i(a,b)}] = \mathbb{E}_{0(0,0)}[T_i] + \mathbb{E}_{U_i}[\tau_{i(a,b)}],$$

where the last expectation means that we start from a uniform point in cluster i and wait to hit $i(a,b)$. Since the graph is transitive, it follows that for all clusters i and all a, b

$$\mathbb{E}_{U_i}[\tau_{i(a,b)}] = z. \quad (4.3)$$

We now define the process Y to be the number of the cluster we are at. More precisely, $Y_t = i$ if and only if $X_t = i(a,b)$ for some a, b . It is easy to check that Y is a Markov chain even with respect to the enlarged filtration which at time t also contains the information about X up to time t . The process Y is a walk on \mathbb{Z}_{12} with additional edges. From that it follows that for all a, b we have

$$\mathbb{E}_{0(a,b)}[T_i] = h(i) = \mathbb{E}[0 \rightarrow i]$$

and $h(i)$ satisfies a system of 6 (by symmetry) linear equations, with solution given by

$$\begin{aligned} h(6) &= 16, \quad h(5) = h(7) = 16, \quad h(4) = h(8) = 15, \\ h(3) &= h(9) = 13, \quad h(2) = h(10) = 13, \quad h(1) = h(11) = 10. \end{aligned} \quad (4.4)$$

Putting everything together, we deduce that for all $i \neq 3, 9$

$$\mathbb{E}_{0(0,0)}[\tau_{i(a,b)}] = \mathbb{E}[0 \rightarrow i] + \mathbb{E}_{U_i}[\tau_{i(a,b)}] = h(i) + z. \quad (4.5)$$

From (4.4), we obtain that

$$\mathbb{E}_{0(0,0)}[\tau_{6(1,1)}] = \max_{\substack{i \neq 3,9 \\ a,b \in \mathbb{Z}_2}} \mathbb{E}_{0(0,0)}[\tau_{i(a,b)}] \quad (4.6)$$

and it remains to show that

$$\mathbb{E}_{0(0,0)}[\tau_{6(1,1)}] \geq \max_{\substack{i=3,9 \\ a,b \in \mathbb{Z}_2}} \mathbb{E}_{0(0,0)}[\tau_{i(a,b)}]. \quad (4.7)$$

Let T be the first time that we hit cluster 3 without using the long edge $0 \rightarrow 3$ directly. It then follows that at time T the position in cluster 3 is uniform. Hence we have

$$\mathbb{E}_{0(0,0)}[\tau_{3(a,b)}] \leq \mathbb{E}[T] + \mathbb{E}_{U_3}[\tau_{3(a,b)}] = \mathbb{E}[T] + z.$$

In view of (4.5) it thus suffices to show

$$\mathbb{E}[T] < \mathbb{E}[0 \rightarrow 6] = 16. \quad (4.8)$$

Let X be the first time that the walk is off the “shuttle” $0 \rightarrow 3$. Then X has the geometric distribution $\mathbb{P}(X = i) = qp^{i-1}$ with $p = 1/6$ and $q = 1 - p = 5/6$. We can now write

$$\mathbb{E}[T] = 1 + \sum_{i=1,3,\dots} \mathbb{P}(X = i) A_1 + \sum_{i=2,4,\dots} \mathbb{P}(X = i) A_2,$$

where A_1 and A_2 are given by

$$\begin{aligned} A_1 &= \frac{2}{5}\mathbb{E}[1 \rightarrow 3] + \frac{2}{5}\mathbb{E}[11 \rightarrow 3] + \frac{1}{5}\mathbb{E}[9 \rightarrow 3] = \frac{72}{5} \\ A_2 &= \frac{2}{5}\mathbb{E}[4 \rightarrow 3] + \frac{2}{5}\mathbb{E}[2 \rightarrow 3] + \frac{1}{5}\mathbb{E}[6 \rightarrow 3] = \frac{53}{5}. \end{aligned}$$

Substituting we deduce

$$\mathbb{E}[T] = \frac{104}{7} < 16,$$

and hence this concludes the proof of the lemma. \square

Proof of Theorem 1.7 (for lazy walk). From Lemma 4.3 we have that the pair that maximizes $\mathbb{E}_x[\tau_y]$ is $x = 0(0,0)$ and $y = 6(1,1)$. (Lemma 4.3 is stated for a non-lazy walk, but the hitting times of the non-lazy and lazy walk are equal up to a factor of 2.) We will now prove that if the moving target stays at position $5(1,1)$ for 2 time steps and then moves to $6(1,1)$, then the expected hitting time is larger than $\mathbb{E}_{0(0,0)}[6(1,1)]$.

We write $\tau_{5 \rightarrow 6}$ for the time to hit the moving target. Then notice that $\tau_{5 \rightarrow 6} - \tau_{6(1,1)}$ is non-zero if we hit 6 at time 1 or 2. We thus have

$$\mathbb{E}_{0(0,0)}[\tau_{5 \rightarrow 6} - \tau_{6(1,1)}] \geq \mathbb{P}_{0(0,0)}(\tau_{6(1,1)} \leq 2) \geq c > 0$$

and this concludes the proof of the theorem for a lazy walk. \square

Proof of Theorem 1.7 (*for continuous time walk*). Consider the graph $G_{7,20}$. Solving the system of expected hitting times and arguing in exactly the same way as in the proof of Lemma 4.3 we get that

$$\mathbb{E}_{0(0,0)}[\tau_{10(1,1)}] = \max_{x,y} \mathbb{E}_x[\tau_y].$$

We now describe a strategy for the moving particle that achieves bigger expected hitting time. Suppose that $f(t) = 8(1,1)$ when $t \leq \varepsilon$ and $f(t) = 10(1,1)$ for $t > \varepsilon$, where $\varepsilon > 0$ will be determined.

Note that $\tau_f - \tau_{10(1,1)}$ is nonzero if and only if $\tau_{10(1,1)} < \varepsilon$ or $\tau_f < \varepsilon$. To simplify notation we write 0 instead of $0(0,0)$ and τ_{10} instead of $\tau_{10(1,1)}$. We now have

$$\begin{aligned} \mathbb{E}_0[\tau_f - \tau_{10}] &= \mathbb{E}_0[(\tau_f - \tau_{10})\mathbf{1}(\tau_f < \varepsilon \text{ or } \tau_{10} < \varepsilon)] \\ &= \mathbb{E}_0[(\tau_f - \tau_{10})\mathbf{1}(\tau_f < \varepsilon)] + \mathbb{E}_0[(\tau_f - \tau_{10})\mathbf{1}(\tau_{10} < \varepsilon, \tau_f > \varepsilon)]. \end{aligned} \quad (4.9)$$

We look at each of these two terms separately. For the first one we get

$$\mathbb{E}_0[(\tau_f - \tau_{10})\mathbf{1}(\tau_f < \varepsilon)] \geq \mathbb{E}_0[\tau_f - \tau_{10} \mid \tau_f < \varepsilon, \tau_f < \tau_{10}] \mathbb{P}_0(\tau_f < \varepsilon, \tau_f < \tau_{10}). \quad (4.10)$$

By the definition of τ_f we have $\{\tau_f < \varepsilon\} = \{\tau_{8(1,1)} < \varepsilon\}$. We now describe an equivalent way of viewing the continuous time chain. To every edge adjacent to a vertex x we assign an exponential clock of parameter $1/d(x)$, where $d(x)$ is the degree of x . Then the Markov chain crosses the edge of the first exponential clock that rings. In order to hit $8(1,1)$ before time ε at least four exponential clocks of a constant parameter should have rung. Thus

$$\mathbb{P}_0(\tau_f < \varepsilon, \tau_{10} > \tau_f) \leq c\varepsilon^4.$$

It is easy to see that there exists a constant c' independent of ε so that

$$\mathbb{E}_0[\tau_{10} - \tau_f \mid \tau_f < \varepsilon, \tau_{10} > \tau_f] \leq c'.$$

Indeed, this expectation can be bounded from above by the commute time between $8(1,1)$ and $10(1,1)$ which is at most twice the distance between $8(1,1)$ and $10(1,1)$ times the total number of edges of $G_{7,20}$. Therefore plugging these estimates in (4.10) we obtain for a positive constant c_1

$$\mathbb{E}_0[(\tau_f - \tau_{10})\mathbf{1}(\tau_f < \varepsilon)] \geq -c_1\varepsilon^4. \quad (4.11)$$

For the second term of (4.9) we have

$$\mathbb{E}_0[(\tau_f - \tau_{10})\mathbf{1}(\tau_{10} < \varepsilon, \tau_f > \varepsilon)] \geq \mathbb{E}_0[(\tau_f - \tau_{10})\mathbf{1}(\tau_{10} < \varepsilon/2, \tau_f > \varepsilon)] \geq \frac{\varepsilon}{2} \mathbb{P}_0(\tau_{10} < \varepsilon/2, \tau_f > \varepsilon)$$

and arguing as above we obtain

$$\mathbb{P}_0(\tau_{10} < \varepsilon/2, \tau_f > \varepsilon) \asymp \varepsilon^2.$$

Putting all these estimates together we deduce

$$\mathbb{E}_0[\tau_f - \tau_{10}] \geq c_2\varepsilon^3 - c_1\varepsilon^4,$$

which can be made strictly positive by choosing $\varepsilon > 0$ sufficiently small and this completes the proof of the theorem. □

Acknowledgements

We are grateful to Guy Kindler for proposing the question that led to this work. We thank Daniel Ahlberg, Almut Burchard, Yuval Peres, Richard Pymar and Alexandre Stauffer for useful discussions. We also thank MSRI, Berkeley, for its hospitality.

References

- [1] David Aldous and J. Fill. *Reversible Markov Chains and Random Walks on Graphs*. In preparation, <http://www.stat.berkeley.edu/~aldous/RWG/book.html>.
- [2] A. Burchard and M. Schmuckenschläger. Comparison theorems for exit times. *Geom. Funct. Anal.*, 11(4):651–692, 2001.
- [3] S. Griffiths, R. J. Kang, R. Imbuzeiro Oliveira, and V. Patel. Tight inequalities among set hitting times in Markov chains. *ArXiv e-prints*, August 2012.
- [4] R. Imbuzeiro Oliveira. Mixing and hitting times for finite Markov chains. *ArXiv e-prints*, August 2011.
- [5] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.
- [6] Roberto Imbuzeiro Oliveira. On the coalescence time of reversible random walks. *Trans. Amer. Math. Soc.*, 364(4):2109–2128, 2012.
- [7] Y. Peres and P. Sousi. An isoperimetric inequality for the Wiener sausage. to appear in *GAFA*.
- [8] Y. Peres and P. Sousi. Mixing times are hitting times of large sets. *ArXiv e-prints*, July 2011.